

Accepted Manuscript

Development of non-linear models predicting daily fine particle concentrations using aerosol optical depth retrievals and ground-based measurements at a municipality in the Brazilian Amazon region

Karen dos Santos Gonçalves, Mirko S. Winkler, Paulo Roberto Benchimol Barbosa, Kees de Hoogh, Paulo Eduardo Artaxo Netto, Sandra de Souza Hacon, Christian Schindler, Nino Künzli

PII: S1352-2310(18)30219-X

DOI: [10.1016/j.atmosenv.2018.03.057](https://doi.org/10.1016/j.atmosenv.2018.03.057)

Reference: AEA 15927

To appear in: *Atmospheric Environment*

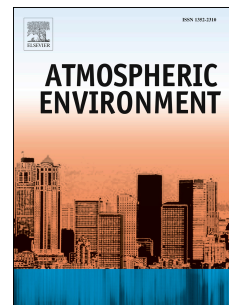
Received Date: 19 June 2017

Revised Date: 19 March 2018

Accepted Date: 28 March 2018

Please cite this article as: Gonçalves, K.d.S., Winkler, M.S., Benchimol Barbosa, P.R., de Hoogh, K., Artaxo Netto, P.E., de Souza Hacon, S., Schindler, C., Künzli, N., Development of non-linear models predicting daily fine particle concentrations using aerosol optical depth retrievals and ground-based measurements at a municipality in the Brazilian Amazon region, *Atmospheric Environment* (2018), doi: 10.1016/j.atmosenv.2018.03.057.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



1 **Development of non- linear models**
2 **predicting daily fine particle concentrations**
3 **using aerosol optical depth retrievals and**
4 **ground-based measurements at a**
5 **municipality in the Brazilian Amazon region**

6 Karen dos Santos Gonçalves^{*,†,‡,§}, Mirko S Winkler^{†,‡}, Paulo Roberto Benchimol
7 Barbosa^{||}, Kees de Hoogh^{†,‡}, Paulo Eduardo Artaxo Netto[⊥], Sandra de Souza Hacon[§],
8 Christian Schindler^{†,‡}, Nino Künzli^{†,‡}

9

10 AUTHORS ADRESS:

11 [†] Swiss Tropical and Public Health Institute, Basel, Switzerland;

12 [‡] University of Basel, Basel, Switzerland;

13 [§] National School of Public Health Sergio Arouca, Oswaldo Cruz Foundation –
14 ENSP/FIOCRUZ, Rio de Janeiro, Brazil;

15 ^{||} Clinical Coordination of Pedro Ernesto University Hospital, Rio de Janeiro State
16 University – HUPE/UERJ, Rio de Janeiro, Brazil;

17 [⊥] Physics Institute, University of São Paulo – IFUSP/USP, São Paulo, Brazil

18

19

20 KEYWORDS: Aerosol Optical Depth; Particulate matter; Air pollution; Forest fire;
21 Validation approach; Brazilian Amazon Region.

22

23 **ABSTRACT**

24

25 Epidemiological studies generally use particulate matter measurements with diameter
26 less $2.5\mu\text{m}$ ($\text{PM}_{2.5}$) from monitoring networks. Satellite aerosol optical depth (AOD)
27 data has considerable potential in predicting $\text{PM}_{2.5}$ concentrations, and thus provides an
28 alternative method for producing knowledge regarding the level of pollution and its
29 health impact in areas where no ground $\text{PM}_{2.5}$ measurements are available. This is the
30 case in the Brazilian Amazon rainforest region where forest fires are frequent sources of
31 high pollution. In this study, we applied a non-linear model for predicting $\text{PM}_{2.5}$
32 concentration from AOD retrievals using interaction terms between average
33 temperature, relative humidity, sine, cosine of date in a period of 365,25 days and the
34 square of the lagged relative residual. Regression performance statistics were tested
35 comparing the goodness of fit and R^2 based on results from linear regression and non-
36 linear regression for six different models. The regression results for non-linear
37 prediction showed the best performance, explaining on average 82% of the daily $\text{PM}_{2.5}$
38 concentrations when considering the whole period studied. In the context of Amazonia,
39 it was the first study predicting $\text{PM}_{2.5}$ concentrations using the latest high-resolution
40 AOD products also in combination with the testing of a non-linear model performance.
41 Our results permitted a reliable prediction considering the AOD- $\text{PM}_{2.5}$ relationship and
42 set the basis for further investigations on air pollution impacts in the complex context of
43 Brazilian Amazon Region.

44

45 **1. INTRODUCTION**

46

47 In spite of the efforts to improve air quality during the past decades, levels of air
48 pollution experienced by human populations continue to cause a large burden of
49 disease.^{1,2,3} Atmospheric aerosols and particulate matter that are breathable ($< 2.5 \mu\text{m}$
50 diameter = $\text{PM}_{2.5}$) and inhalable ($< 10 \mu\text{m}$ = PM_{10}), generated from natural and
51 anthropogenic emission sources present known effects for a number of causes of death,
52 particularly the increase in cardio-respiratory diseases in areas with high
53 concentrations.^{4,5}

54

55 Intensive and indiscriminate occurrence of forest fire has become a serious
56 environmental problem in Brazil, affecting ecosystems' balance and human health with
57 consequences at the local, regional and global level.^{6,7} Brazilian Amazon region has
geographic and environmental circumstances that are distinct from other world regions.

58 For this reason, the occurrence of fire and emissions of PM_{2.5} exposes every year
59 increasingly large portions of vulnerable populations.^{8,9}

60 To understand the association between PM_{2.5} and effects on human health,
61 epidemiological studies have employed PM_{2.5} measurements from monitoring sites.
62 However, due to cost and lack of appropriate infrastructure, especially in rural and
63 remote areas, no fixed site PM_{2.5} measurements are available in many regions of Brazil.
64 This is a major limitation for estimating exposure to PM_{2.5} and assessing health impacts
65 associated with forest fires as one of its major source.^{10,11,12,13,14,15}

66 An alternative approach to estimate the air quality in areas without direct PM_{2.5}
67 measurements is by means of satellite remote sensing using aerosols optical depth
68 (AOD). AOD is an electromagnetic radiation measure and reflects the integrated
69 number of particles at a given wavelength. It is an important satellite-retrieved property
70 for predicting the PM_{2.5} concentrations due repeated observations of the atmosphere and
71 its extensive spatial coverage.¹⁶ The AOD has been successfully used in statistical
72 models for estimating PM_{2.5} levels. As shown by previous studies, parameters such as
73 local meteorology and land use information influence the relationship between AOD
74 and daily PM_{2.5} concentrations, which need to be considered as additional
75 predictors.^{10,17,18,19,20,21,22,23,24}

76 Traditionally, the health exposure studies have used the standard MODIS (Moderate
77 Resolution Imaging Spectroradiometer) AOD product of the “Dark Target” algorithm
78 published by Levy et al. (2007, 2010), which has a resolution of 10 x 10 km². Later,
79 Remer et al. (2013, 2005) described AOD algorithm applying a higher resolution of 3 X
80 3 Km².^{25,26,27,28}

81 Concerning the applicability of the statistical methods for predicting PM_{2.5}
82 concentration using AOD retrievals, de Hoogh K et al (2017)²⁹ used a higher spatial
83 resolution for modelling daily PM_{2.5} concentrations across Switzerland during the
84 period between 2003 to 2013. Their models result explained on average 73% of the total
85 ,71% of the spatial and 75% of the temporal variation (all cross validated) in measured
86 PM_{2.5} concentrations. Kloog Itai et al. (2017)³⁰ described a new hybrid spatio-temporal
87 model for estimating daily PM_{2.5} concentrations across northeastern USA using high
88 resolution AOD data. Their results showed a high predictive accuracy at high spatial
89 resolutions using a mixed model regressing PM_{2.5} measurements with an excellent
90 model performance (R²=0.88).

91

92 These recent studies still have the challenge of reducing exposure error, although
93 shows better fits than previous models. In spite our model showed a good performance,
94 it is important to reproduce it in another region with different meteorological and
95 geographical patterns. Our model can be applied to other sites if site-specific AOD and
96 meteorological data are available, to be inserted into the prediction equation. The lagged
97 relative residual added as a further predictor variable it is cautious strategy to remove
98 the serial autocorrelation and to further improve the model. As another important
99 challenge is that AOD data availability is much greater in the dry seasons compared to
100 the rainy period. This is mostly due to heavily clouded days which results in missing
101 AOD data. This non-random lack of AOD readings could negatively affect predictive
102 performance. Also, treating large areas, such as Brazilian Amazon region, can add
103 additional selection bias since there may be meteorological variations in the daily
104 calibration between PM_{2.5} and AOD.³⁰

105 In this paper we developed a non-linear model predicting daily fine particle
106 concentrations using AOD retrievals at 3 x 3 km resolution and ground-based
107 measurements at a municipality of Porto Velho, Brazil during the period between 2009
108 to 2011. For Brazilian Amazon region, it is the first study to develop this approach
109 considering a non-linear model predicting PM_{2.5} concentrations. This study assessment
110 is part of an investigation that aims at analysing the impact of PM_{2.5} exposure on
111 cardiovascular disease in Porto Velho.

112

113 2. MATERIAL AND METHODS

114 2.1. Ground-level PM_{2.5} data

115 Daily averages were derived during the period from 25 September 2009 to 21
116 October 2011 with a total of 757 days. Over the study period, PM_{2.5} concentrations were
117 measured for 24h at one air quality monitoring station in Porto Velho municipality,
118 which was implanted in partnership between Institute of Physics at University of São
119 Paulo (USP), University of Rondônia (UNIR), Environmental Biogeochemistry
120 Laboratory Wolfgang H. Pfeiffer and National School of Public Health-Oswaldo Cruz
121 Foundation (FIOCRUZ) in Brazil. The PM_{2.5} monitor is located at 15 km north of to the
122 centre of urban area (**Figure 1**). Porto Velho municipality is the third capital in the
123 Brazilian Amazon region with 67 districts within the urban area. With an area of 34,096
124 km² Porto Velho has a population of 503,000 inhabitants according to *Brazilian*
125 *Institute of Geography and Statistics (IBGE, Census 2010)*. PM_{2.5} measurements were

126 collected by means of a stacked filter unit (SFU) and were analysed gravimetrically
127 according to the World Health Organization Air Quality Guidelines for particulate
128 matter, ozone, nitrogen dioxide and sulfur dioxide (WHO, 2005)³¹.

129 This methodology involves the sampling site (8.69° S, 63.87° W) located in a region
130 with large land use changes and associated regional biomass burning. The SFU (Stacked
131 Filter Unit) type samplers and the analysis follows routine gravimetric techniques³³. In
132 addition, trace elements and ionic compounds are collected, allowing for future
133 analyses.

134 There is an AFG sampler, which collects aerosols for elemental PIXE and black
135 carbon analyses on the roof of the shelter, 24 hours sampling. The collection of aerosol
136 particles using filters is a simple and very common method for sampling aerosol
137 particles. Filters allow elemental and ionic analysis through a series of measurement
138 techniques. The sampler collects fine and coarse particles and contains an inlet that
139 allows the entry of particles in the range of $2 < D_p < 10 \mu\text{m}$. The filters are polycarbonate,
140 having a diameter of 47mm and are arranged in series. In the first step the particles of
141 the coarse fraction are retained using Nucleopore filters with pores of $8 \mu\text{m}$ in diameter,
142 in the second stage, they are the fine particles that are retained using the filter
143 Nucleopore with pores of $0.4 \mu\text{m}$. The samples collected with the AFG sampler was
144 used to determine the mass of the aerosols by means of gravimetric analysis, the
145 concentration of black carbon and to quantify the elemental concentration of the
146 material deposited in the filters.

147

148 **2.2. MODIS 3 km AOD retrieval**

149 The Moderate Resolution Imaging Spectroradiometer (MODIS) is a key instrument
150 aboard the Terra and Aqua satellites of the National Aeronautics and Space
151 Administration (NASA) and has been in operation since 1999 and 2002, respectively.
152 While Terra passes the equator in the morning, from north to south, Aqua passes the
153 equator from south to north in the afternoon. These satellites were used to retrieve AOD
154 aerosol products with a 3 km resolution (MOD04_3K and MYD04_3K), operating at an
155 altitude of approximately 700 km (<http://modis-atmos.gsfc.nasa.gov/>). In the Collection
156 6, Level 2 aerosol products, the most recent 3 km AOD dark target retrieval algorithm is
157 similar to the 10 km standard product (Collection 5, Level 2) and has three different
158 wavelength channels of 0.47, 0.66 and $2.12 \mu\text{m}$ employed for AOD retrieval over land.
159 The other channels are used for screening procedures (e.g., coverage of cloud, snow and

160 ice).^{25,26,32} More details on the retrieval of MODIS satellite aerosol data have previously
 161 been published by Remer et al. (2013, 2005)^{25,28} and Levy et al. (2007, 2010)^{26,27}. For
 162 the AOD daily averages, we used the algorithm retrieval in MATLAB (version 2015a,
 163 MathWorks) and the software ArcGIS (version 10, ESRI) to create 820 grid cells of 3 x
 164 3-km covering the study area for spatial analyses.

165

166 **2.3. Statistical model and validation**

167 In this study we considered five different types of prediction models of PM_{2.5}
 168 concentrations from AOD retrievals. They were all of the form
 169 $PM_{2.5} = \exp(\text{linear predictor})$, with the linear predictor involving terms composed of
 170 AOD and other influencing factors. The advantage of such a model over one for log-
 171 transformed outcome data is that it provides estimates of mean exposure levels while
 172 estimates of geometric mean levels are obtained when modelling log-transformed
 173 outcome data and then exponentiating the resulting predictions (which are on the log-
 174 scale). In Model 1, we took the linear predictor to be a cubic polynomial in AOD with
 175 time-independent coefficients:

176

(Model 1)

177

$$PM_{2.5} = \exp(\alpha + \beta_1(AOD) + \beta_2(AOD)^2 + \beta_3(AOD)^3)$$

178

179 In a next step, we considered the coefficients of Model 1 to be polynomials of second
 180 degree in average temperature (TEMP) and relative humidity (RH). Thus, each of the
 181 coefficients was assumed to be of the form $\gamma_0 + \gamma_1 \cdot \text{TEMP} + \gamma_2 \cdot \text{RH} + \gamma_3 \cdot \text{TEMP}^2 + \gamma_4 \cdot \text{RH}^2$
 182 $+ \gamma_5 \cdot \text{TEMP} \cdot \text{RH}$. Multiplying these polynomials with AOD, AOD² and AOD³ and the
 183 intercept, respectively, provided:

184

(Model 2)

185

186

187

188

189

$$PM_{2.5} = \text{Model 1} \cdot \exp(\beta_4(AOD \cdot \text{TEMP}) + \beta_5(AOD \cdot \text{TEMP}^2) + \beta_6(AOD \cdot \text{RH}) + \\ \beta_7(AOD \cdot \text{RH}^2) + \beta_8(AOD \cdot \text{TEMP} \cdot \text{RH}) + \beta_9(AOD^2 \cdot \text{TEMP}) + \beta_{10}(AOD^2 \cdot \text{TEMP}^2) + \\ \beta_{11}(AOD^2 \cdot \text{RH}) + \beta_{12}(AOD^2 \cdot \text{RH}^2) + \beta_{13}(AOD^2 \cdot \text{TEMP} \cdot \text{RH}) + \beta_{14}(AOD^3 \cdot \text{TEMP}) + \\ \beta_{15}(AOD^3 \cdot \text{TEMP}^2) + \beta_{16}(AOD^3 \cdot \text{RH}) + \beta_{17}(AOD^3 \cdot \text{RH}^2) + \beta_{18}(AOD^3 \cdot \text{TEMP} \cdot \text{RH}))$$

190

191

192

193

In an attempt to further improve the model, we added interaction terms between
 AOD, AOD² and AOD³ with rainy season (Model 3) and tested interactions of the three
 AOD-terms with sine and cosine functions of date with a period of 365.25 days in the
 Model 4.

194 In the last step, we included the lagged relative residual and its square as additional
 195 predictor variables (Model 5). This was to reduce serial autocorrelation. The final model
 196 obtained after some backward elimination steps was of the form:

197 (Final model)

$$\begin{aligned}
 198 \text{PM}_{2.5} = & \exp(\alpha + \beta_1 \cdot \text{AOD} + \beta_2 \cdot \text{AOD}^2) + \beta_3 \cdot \text{AOD}^3 + \beta_4 \cdot (\text{AOD} \cdot \text{TEMP}) + \beta_5 (\text{AOD} \cdot \text{RH}) \\
 199 & + \beta_6 (\text{AOD} \cdot \text{TEMP}^2) + \beta_7 (\text{AOD} \cdot \text{RH}^2) + \beta_8 (\text{AOD} \cdot \cos_days) + \beta_9 (\text{AOD} \cdot \sin_days) + \\
 200 & \beta_{10} (\text{AOD}^2 \cdot \cos_days) + \beta_{11} (\text{AOD}^2 \cdot \sin_days) + \beta_{12} (\text{AOD}^3 \cdot \cos_days) + \\
 201 & \beta_{13} (\text{AOD}^3 \cdot \sin_days) + \beta_{14} (\text{residual}) + \beta_{15} (\text{residual})^2
 \end{aligned}$$

202

203 In these equations, $\text{PM}_{2.5}$ denotes the predicted concentrations, where *exp* is the
 204 exponential function; *cos_days* and *sin_days* denote the cosine and sine terms of date
 205 with a period of 365,25 days; *residual* denotes the lagged relative residual. Relative
 206 residuals were defined as ratio between residuals and predicted values. Model
 207 performance was evaluated by comparing the predictions with the ground measurements
 208 using the adjusted coefficient of determination (R^2 adj), residual standard deviation
 209 (RMSE), Akaike's information criterion (AIC), and partial autocorrelation of residuals
 210 by lags. High values of adjusted R squared suggest that MODIS AOD data can be used
 211 to estimate ambient concentrations. Furthermore, we calculated mean, standard
 212 deviation and maximum / minimum values to summarize the descriptive statistics of our
 213 sample for the whole period, the dry season (months June to November, when forest
 214 fires occur in Brazilian Amazon region) and the rainy season (months from December
 215 to May).

216 Daily meteorological data on average temperature and relative humidity were
 217 obtained for the period 25 September 2009 to 21 October 2011 from the monitoring
 218 station of INMET (*National Meteorology Institute*) in Porto Velho. The information is
 219 publicly available on the website of the institute (www.inmet.gov.br).

220 The spatial distribution of the 3x3km resolution MODIS AOD average during the
 221 study period was derived by spatial interpolation using the inverse distance weighting
 222 (IDW). We present the results for all states in Brazil and for our study area. It is
 223 important to highlight that all regression results were presented with the original AOD
 224 and $\text{PM}_{2.5}$ datasets. The software R (version 3.1.3) was used for statistical analyses.

225

226 2. RESULTS

227 3.1. Descriptive statistics

228 **Table 1** shows the descriptive statistics of daily measured PM_{2.5}, values of AOD,
229 relative humidity, average temperature and precipitation from 25 September 2009 to 21
230 October 2011, as well as for the dry and rainy seasons. Average daily level of PM_{2.5}
231 from the ground-level monitor was 11µg/m³ with a standard deviation (SD) of
232 ±20µg/m³ over the three years studied. Of note, the analysis of the data for 2010, the
233 year when one of the most extreme dry seasons in Brazilian Amazon region occurred,
234 revealed an annual-mean of 36µg/m³ (±46µg/m³ SD).

235 Considering the differences between seasons, the maximum daily value was
236 exceptionally high (164µg/m³) during the dry seasons of the period studied compared to
237 27µg/m³ in the rainy seasons.

238 Over the entire study period, the daily AOD values observed varied from 0.03 to
239 2.19. On average, 649 AOD values were retrieved per grid cell which corresponds to
240 86% of the entire study period of 757 days.

241 All the meteorological variables such as relative humidity, average temperature and
242 precipitation were consistent with the climatic patterns expected for the Brazilian
243 Amazon region and thus support the analysis in the regression models.

244

245 **3.2. Non-linear prediction models**

246 To test the performance of the five regressions models we use a total of 649 valid
247 days for the model fitting. The comparisons between the models analysed and
248 parameters estimated are shown in **Table 2 and 3**.

249 Model 1 shows an adj R² of 0.54, RMSE of 13.59µg/m³ and AIC of 5234.3 for the
250 whole period. Model 2 including interactions between AOD, AOD² and AOD³ and
251 linear and quadratic terms in temperature and relative humidity provided a better fit (R²
252 = 0.67). After adding interactions between the three AOD-terms and rain the fit only
253 slightly improved (R²=0.70). In Model 4 we excluded the rain term and added
254 interactions of the three AOD-terms with sine and cosine of date with a period of
255 365.25 days. This model performed considerably better (R²=0.77; RMSE=9.59µg/m³;
256 AIC=4803.1).

257 After adding the lagged relative residual and its square as additional predictors
258 (Model 5) the adjusted R² further increased to 0.82 (RMSE=8.60µg/m³; AIC=4797.6).
259 This means that this non-linear prediction model explains 82% of the variance of daily
260 PM_{2.5} concentrations in combination with meteorological and seasonal variables. The
261 introduction of these two terms also led to a drastic reduction in residual

262 autocorrelation, in that lag1-autocorrelation of residuals was no longer significant
263 (**Figure 2**). As visualized in **Figure 2**, the time series of predicted $PM_{2.5}$ concentrations
264 follows a very similar pattern as the measured $PM_{2.5}$ confirming the high performance
265 of the prediction models. The period from mid-July to end of October 2010 – a dry
266 period with plumes of biomass burning – is characterized by very high AOD values,
267 reaching peaks 50-100 times above the typical values observed before and after this
268 period. The comparisons between the measured and predicted $PM_{2.5}$ concentrations for
269 Model 1 and Model 5 are illustrated in **Figure 3**.

270 The Spatial distribution of $PM_{2.5}$ predicted concentration over the basin during
271 different seasons for all Brazilian states are shown in **Figure 4**. The highest predicted
272 $PM_{2.5}$ concentrations were observed in the Brazilian Amazon region during the forest
273 fires season (months between September, October and November). In our study area,
274 $PM_{2.5}$ averages reached $44\mu g/m^3$ in the urban area of Porto Velho, and $54\mu g/m^3$ across
275 the Rondônia state during the forest fires between 2009 and 2011. Information about the
276 distribution of $PM_{2.5}$ within the urban districts of Porto Velho and the relation with the
277 health data will be presented in a separate manuscript about the impacts of $PM_{2.5}$ on
278 human health in Brazilian Amazon region.

279

280 3. DISCUSSION

281 The results of our final non-linear prediction model for $PM_{2.5}$ showed a good
282 performance, explaining on average 82% of the variance in measured $PM_{2.5}$
283 concentrations during the period studied. This result is similar and in accordance with
284 the findings presented by Lee et al (2011)¹⁰ and Xie et al (2015)²⁴, who showed
285 prediction models that explained 92% and 82% of the variance in $PM_{2.5}$ concentrations
286 in the North-eastern, US and in Beijing, China, respectively. Our model has the
287 advantage that it does not produce negative predictions and fits the mean of the data as a
288 function of the predictor variables. Moreover, by including the lagged relative residual
289 and its square as additional predictor variables it was possible to remove the significant
290 lag1-autocorrelation, and to further improve the model fit.

291 Observing the temporal distribution of predicted and measured $PM_{2.5}$ concentrations
292 it is important to highlight the enormous peak of $PM_{2.5}$ observed between days 300 and
293 400 during the dry periods in our study area with the maximum daily value of
294 $164\mu g/m^3$. This value more than 6.5 times higher than the daily mean guideline value
295 proposed by WHO to protect public health ($25\mu g/m^3$). During the dry season, this value

296 was exceeded on X% of all days. As a consequence, the long-term mean concentration
297 during the dry period was 2 times above the WHO annual mean guideline value, set at
298 $10\mu\text{g}/\text{m}^3$. Currently, these values were adopted in only a few countries as legally
299 binding targets, thus, policy makers accept major impacts on morbidity and mortality.³⁴
300 On the other hand, during the rainy seasons concentrations were low ($2\mu\text{g}/\text{m}^3$; $\pm 3\mu\text{g}/\text{m}^3$
301 SD) and fully in line with both, the daily and annual targets proposed by WHO. This
302 confirms the dominant role of fires as source of ambient air pollution in the Amazon
303 region. This result highlight the importance to set limits for $\text{PM}_{2.5}$ in the Brazil Air
304 quality Standards defined by the National Environmental Agency (CONAMA) that
305 currently set limits only for PM_{10} .³⁴ Our data provide unique input to evaluate whether
306 the CONAMA may add to the Brazil Air Q S limits for $\text{PM}_{2.5}$ rather than PM_{10} alone.
307 This will be particularly worth if sources and spatio-temporal patterns of the two
308 markers of air pollution largely vary across Brazil.

309 The non-linear prediction model demonstrated a high performance in predicting the
310 daily $\text{PM}_{2.5}$ concentrations. However, some limitations, such as cloud properties and
311 uncertainties need to be mentioned. The use of only one air monitoring station for the
312 development and evaluation of the model is a major limitation of this study. However,
313 although this limits our ability to draw firm conclusions about the applicability of the
314 model across Brazil, it does provide a valid approach to predict $\text{PM}_{2.5}$ across our main
315 study area. $\text{PM}_{2.5}$ is spatially rather homogenously distributed, thus extrapolation of the
316 model from the measurement site to the adjacent urban area of Porto Velho is expected
317 to be reliable. The model gives also good indications of possible hot spots of pollution
318 where it may be worth installing additional monitors operating continuously or at least
319 during dry seasons. With the use of mobile stations, one could characterize the spatial
320 pattern of air pollution across a larger area with only one or a few monitors while using
321 the current central monitoring station as a reference point to understand the temporal
322 variation. The installation of air quality monitoring networks is an important step for the
323 future evaluation of progress in clean air management and the assessment of its health
324 impact.

325 The predictions were based on different spatial scales which may be a source of
326 uncertainties. The AOD satellite data is based on a grid cell with 3km resolution while
327 $\text{PM}_{2.5}$ ground-level is measured at a fixed point. It is also important to highlight the
328 complex relationship between AOD and $\text{PM}_{2.5}$ due the nature of the forest aerosol in

329 Brazilian Amazon. Other important predicting factors such as wind speed, atmosphere
330 or physical-chemical components were not analysed in this study.

331 Despite the uncertainties, this study is the first to predict $PM_{2.5}$ concentrations using
332 a non-linear prediction model and the higher-resolution MODIS AOD products in Porto
333 Velho. By modelling the AOD- $PM_{2.5}$ relationship in a time-dependent manner reflecting
334 seasonal fluctuations and influences of temperature and relative humidity we were able
335 to develop a prediction model for $PM_{2.5}$ with good fitting properties.

336 Our model also can be applied to other sites of the region if site-specific AOD- and
337 meteorological data are available, to be inserted into the prediction equation. On the
338 other hand, as measured $PM_{2.5}$ data are only available for the reference station, the
339 relative lag1-residuals of $PM_{2.5}$ at the reference site would have to be used for all other
340 sites. About the applicability of the model to other sites, the Model 5 (not involving
341 lagged residuals) could be applied to any other site of the Amazon region, as AOD-
342 values and estimates of meteorological parameters can be obtained for any such site
343 based on satellite data. On the other hand, model 6 involves an input variable which is
344 only available at the reference site and whose values would therefore have to be used at
345 all other sites. A cautious strategy might therefore consist in using both model 5 and 6
346 when conducting time series analyses of deaths and hospital admissions.

347

348 4. CONCLUSIONS

349 Satellite data has an important potential for the spatio-temporal prediction of $PM_{2.5}$
350 concentrations. It offers an alternative method to describe the impacts of forest fires on
351 air quality and to assess the related health effects in the Brazilian Amazon region. Our
352 method provides valuable inputs on how to strengthen and optimize the $PM_{2.5}$
353 monitoring networks with well-placed complementary measurement sites.

354 This is needed to understand the impact of air pollution in Brazil and to demonstrate
355 the improvements of air quality and the related health benefits due to the adoption of
356 clean air policies.

357

358 AUTHOR INFORMATION

359 **Corresponding author:**

360 ***Postal address:** ENSP/FIOCRUZ: Rua Leopoldo Bulhões, 1480, CEP: 21041-210,

361 Manguinhos, Rio de Janeiro, Brazil; **E-mail:** karengoncalves@gmail.com

362

363 **Note:** The authors declare no conflict of interest.

364

365

366 **ACKNOWLEDGEMENT**

367

368 This work was supported by the CAPES foundation [grant number
369 88881.068027/2014-01] from Brazil. Special thanks to Glauber G.C. Silva
370 (INPE/Brazil), who was always kind and available to help with MATLAB scripting and
371 technical solutions. This manuscript is part of a larger project on *Cardiovascular*
372 *diseases and the exposure to forest fires in Porto Velho municipality, Rondônia state,*
373 *Brazil* that was submitted and accepted by the Research Ethics Committee of the Sergio
374 Arouca National Public Health School (*Comitê de Ética em Pesquisa da Escola*
375 *Nacional de Saúde Pública Sergio Arouca – ENSP / FIOCRUZ*) according to
376 Resolution number 466/2012 from National Research Ethics Council (*Conselho*
377 *Nacional de Pesquisa – CONEP*) under CAAE number 41732615.4.0000.5240.

378

379

380 **REFERENCES**

381 (1) Cohen, AJ; Anderson, HR; Ostro, B; Pandey, KD; Krzyzanowski, M; Künzli, N;
382 Gutschmidt, K; Pope, A; Romieu, I; Samet, JM & Smith, K. The Global Burden of
383 Disease Due to Outdoor Air Pollution. *Journal of Toxicology and Environmental*
384 *Health, Part A*, 2005. 68:1301–1307, doi:10.1080/15287390590936166

385

386 (2) Brauer M, Freedman G, Frostad J, van Donkelaar A, Martin RV, Dentener F, Van
387 Dingenen R, Estep K, Amini H, Apte JS, Balakrishnan K, Barregard L, Broday
388 DM, Feigin V, Ghosh S, Hopke PK, Knibbs LD, Kokubo Y, Liu Y, Ma S,
389 Morawska L, Sangrador JLT, Shaddick G, Anderson HR, Vos T, Forouzanfar MH,
390 Burnett RT, Cohen A. Ambient air pollution exposure estimation for the Global
391 Burden of Disease 2013. *Environmental Science & Technology*. 2015 Nov 23. doi:
392 10.1021/acs.est.5b03709.

393

394 (3) Global Burden of Diseases (GBD) (2010) Institute for Health Metrics and
395 Evaluation 2013. <http://vizhub.healthdata.org/irank/heat.php>

396

397 (4) Brook, RD; Rajagopalan, S; Pope III, CA; Brook, JR; Bhatnagar, A; Diez-Roux,
398 AV; Holguin, F; Hong, Y; Luepker, RV; Mittleman, MA; Peters, A; Siscovick, D;

- 399 Smith, SC; Whitsel, L & Kaufman, JD. Particulate matter air pollution and
400 cardiovascular disease: An update to the scientific statement from the American
401 Heart Association. *Journal AHA. Circulation* 2010. 2331-2378, doi:
402 10.1161/CIR.0b013e3181d8e1;
403
- 404 (5) World Health Organization (WHO) (2014) 7 million premature deaths annually
405 linked to air pollution. [http://www.who.int/mediacentre/news/releases/2014/air-](http://www.who.int/mediacentre/news/releases/2014/air-pollution/en/)
406 [pollution/en/](http://www.who.int/mediacentre/news/releases/2014/air-pollution/en/)
407
- 408 (6) Gonçalves, KS; Castro, HA; Hacon, SS. As queimadas na região amazônica e o
409 adoecimento respiratório. *Rev Ciência e Saúde Coletiva* 2012. 17(6): 1523-1532,
410 doi: <http://dx.doi.org/10.1590/S1413-81232012000600016>;
411
- 412 (7) Becker BK. Geopolítica da Amazônia. *Estudos avançados* 2005; 1(53):71-86.
413
- 414 (8) Fearnside, P.M. Desmatamento na Amazônia brasileira: história, índices e
415 conseqüências. MEGADIVERSIDADE, Vol 1 (1) Julho 2005, Available:
416 [http://philip.inpa.gov.br/publ_livres/2005/Desmatamento%20historia-](http://philip.inpa.gov.br/publ_livres/2005/Desmatamento%20historia-Megadiversidade.pdf)
417 [Megadiversidade.pdf](http://philip.inpa.gov.br/publ_livres/2005/Desmatamento%20historia-Megadiversidade.pdf), Accessed: May/2016;
418
- 419 (9) Goncalves,KS; Siqueira,ASP,Castro, HA; Hacon, SS.Indicator of socio-
420 environmental vulnerability in the Western Amazon. The case of the city of Porto
421 Velho, State of Rondônia, Brazil.*Ciência & Saúde Coletiva*, 19(9):3809-3817, 2014
422 doi: 10.1590/1413-81232014199.14272013
423
- 424 (10) Lee, H. J., Liu, Y., Coull, B.A., Schwartz, J. and Koutrakis, P. A calibration
425 method of MODIS AOD data to predict PM2.5. *Atmos. Chem. Phys.*, 11, 7991–
426 8002, 2011- doi: 10.5194/acp-11-7991-2011
427
- 428 (11) Ruckerl R, Schneider A, Breitner S, Cyrus J, Peters A. Health effects of
429 particulate air pollution: A review of epidemiological evidence. *Inhal Toxicol.*
430 2011Aug; 23(10):555-92, doi: 10.3109/08958378.2011.593587.
431

- 432 (12) Ye X, Wolff R, Yu W, Vaneckova P, Pan X, Tong S. Ambient Temperature and
433 Morbidity: A Review of Epidemiological Evidence. *Environ Health Perspect.* 2011
434 Aug, doi: 10.1289/ehp.1003198.
- 435
- 436 (13) Yi O, Hong YC, Kim H. Seasonal effect of PM(10) concentrations on mortality
437 and morbidity in Seoul, Korea: a temperature-matched case-crossover analysis.
438 *Environ Res.* 2010 Jan;110(1):89-95, doi: 10.1016/j.envres.2009.09.009.
- 439
- 440 (14) Arbex MA, de Souza Conceição GM, Cendon SP, Arbex FF, Lopes AC, Moysés
441 EP, et al. Urban air pollution and chronic obstructive pulmonary disease-related
442 emergency department visits. *J Epidemiol Community Health.* 2009
443 Oct;63(10):777-83, doi: 10.1136/jech.2008.078360 ;
- 444
- 445 (15) McMichael AJ, Wilkinson P, Kovats RS, Patternden S, Hajat S, Armstrong B,
446 et al. International study of temperature, heat and urban mortality: the
447 'ISOTHURM' project. *Int J Epidemiol.* 2008 Oct;37(5):1121-31, doi:
448 10.1093/ije/dyn086
- 449
- 450 (16) Kloog, I.; Chudnovsky, A. A.; Just, A. C.; Nordio, F.; Koutrakis, P.; Coull, B.
451 A.; Lyapustin, A.; Wang, Y.; Schwartz, J., A new hybrid spatio-temporal model for
452 estimating daily multi-year PM 2.5 concentrations across northeastern USA using
453 high resolution aerosol optical depth data. *Atmospheric Environment* **2014**, 95, 581-
454 590, doi:10.1016/j.atmosenv.2014.07.014
- 455
- 456 (17) Liu, Y., Sarnat, J. A., Kilaru, V., Jacob, D. J., and Koutrakis, P.: Estimating
457 ground-level PM2.5 in the eastern United States using satellite remote sensing,
458 *Environ. Sci. Technol.*, 39, 3269–3278 ; 2005.
- 459
- 460 (18) Liu, Y., Franklin, M., Kahn, R., and Koutrakis, P.: Using aerosol optical thickness
461 to predict ground-level PM2.5 concentrations in the St. Louis area: A comparison
462 between MISR and MODIS, *Remote Sens. Environ.*, 107, 33–44, 2007a.
- 463
- 464 (19) Liu, Y., Koutrakis, P., and Kahn, R.: Estimating fine particulate matter component
465 concentrations and size distributions using satellite-retrieved fractional aerosol
466 optical depth: Part 1 – Method development, *J. Air Waste Manag. Assoc.*, 57,
467 1351– 1359, 2007b.

- 468
469 (20) Liu, Y., Koutrakis, P., Kahn, R., Turquety, S., and Yantosca, R. M.: Estimating
470 fine particulate matter component concentrations and size distributions using
471 satellite-retrieved fractional aerosol optical depth: Part 2 – A case study, *J.*
472 *AirWaste Manag. Assoc.*, 57, 1360–1369, 2007c.
- 473
474 (21) Liu, Y., Paciorek, C. J., and Koutrakis, P.: Estimating regional spatial and temporal
475 variability of PM_{2.5} concentrations using satellite data, meteorology, and land use
476 information, *Environ. Health Persp.*, 117, 886–892, 2009
- 477
478 (22) Hoff, R. M. and Christopher, S. A.: Remote sensing of particulate pollution from
479 space: Have we reached the promised land?, *J. Air Waste Manag. Assoc.*, 59, 645–
480 675, 2009, doi: 10.3155/1047-3289.59.6.645;
- 481
482 (23) Liu, Y., Park, R. J., Jacob, D. J., Li, Q. B., Kilaru, V., and Sarnat, J. A.: Mapping
483 annual mean ground-level PM_{2.5} concentrations using Multiangle Imaging
484 Spectroradiometer aerosol optical thickness over the contiguous United States, *J.*
485 *Geophys. Res.*, 109, D22206, doi:10.1029/2004JD005025, 2004.
- 486
487 (24) Xie, Yuanyu; Wang, Yuxuan; Zhang, Kai; Dong, Wenhao; Lv, Baolei and Bai,
488 Yuqi. Daily Estimation of Ground-Level PM_{2.5} Concentrations over Beijing Using
489 3km Resolution MODIS AOD *Environ. Sci. Technol.*, 2015, 49 (20), pp 12280–
490 12288 DOI: 10.1021/acs.est.5b01413
- 491
492 (25) L. A. Remer S. Mattoo, R. C. Levy, and L. A. Munchak. MODIS 3km aerosol
493 product: algorithm and global perspective. *Atmos. Meas. Tech.*, 6, 1829–1844,
494 2013
495 www.atmos-meas-tech.net/6/1829/2013/ doi:10.5194/amt-6-1829-2013
- 496
497 (26) Levy, R. C., Remer, L. A., Mattoo, S., Vermote, E. F., and Kaufman, Y. J.: Second-
498 generation operational algorithm: Retrieval of aerosol properties over land from
499 inversion of Moderate Resolution Imaging Spectroradiometer spectral reflectance,
500 *J. Geophys. Res.*, 112, D13211, doi:10.1029/2006JD007811, 2007.
- 501

- 502 (27) Levy, R. C., Remer, L. A., Kleidman, R. G., Mattoo, S., Ichoku, C., Kahn, R., and
503 Eck, T. F.: Global evaluation of the Collection 5 MODIS dark-target aerosol
504 products over land, *Atmos. Chem. Phys.*, 10, 10399–10420, doi:10.5194/acp-10-
505 10399-2010, 2010.
- 506
- 507 (28) Remer, L. A.; Kaufman, Y. J.; Tanre, D.; Mattoo, S.; Chu, D. A.; Martins, J. V.; Li,
508 R. R.; Ichoku, C.; Levy, R. C.; Kleidman, R. G.; et al. *Environmental Science &*
509 *Technology*. The MODIS aerosol algorithm, products, and validation. *J. Atmos.*
510 *Sci.* 2005, 62 (4), 947–973. doi: 10.1021/acs.est.5b01413;
- 511
- 512 (29) Kees de Hoogh, Harris Héritier, Massimo Stafoggia, Nino Künzli, Itai Kloog,
513 Modelling daily PM_{2.5} concentrations at high spatio-temporal resolution across
514 Switzerland, In *Environmental Pollution*, 2017, , ISSN 0269-7491,
515 <https://doi.org/10.1016/j.envpol.2017.10.025>.
- 516
- 517 (30) Kloog, I.; Nordio, F.; Coull, B. A.; Schwartz, J., Incorporating Local Land Use
518 Regression And Satellite Aerosol Optical Depth In A Hybrid Model Of
519 Spatiotemporal PM_{2.5} Exposures In The Mid-Atlantic States. *Environmental*
520 *Science & Technology* **2012**, 46 (21), 11913-11921. dx.doi.org/10.1021/es302673e
- 521
- 522 (31) World Health Organization (WHO) (2006) Air quality guidelines for particulate
523 matter, ozone, nitrogen dioxide and sulfur dioxide—global update 2005 Geneva:
524 WHO Office for Europe.
525 http://www.euro.who.int/__data/assets/pdf_file/0005/78638/E90038.pdf
- 526
- 527 (32) Munchak, L. A.; Levy, R. C.; Mattoo, S.; Remer, L. A.; Holben, B. N.; Schafer, J.
528 S.; Hostetler, C. A.; Ferrare, R. A. MODIS 3km aerosol product: applications over
529 land in an urban/suburban region. *Atmos. Meas. Tech.* 2013, 6 (7), 1747–1759,
530 <https://doi.org/10.5194/amt-6-1747-2013>
- 531
- 532 (33) Philip K. Hopke , Ying Xie , Taisto Raunemaa , Steven Biegalski , Sheldon
533 Landsberger, Willy Maenhaut , Paulo Artaxo & David Cohen (1997)
534 Characterization of the Gent Stacked Filter Unit PM₁₀ Sampler, *Aerosol Science*

535 and Technology, 27:6, 726-735, DOI: 10.1080/02786829708965507

536

537 (34) Künzli N, Kutlar Joss MK, Gintowt E (2015) Global standards for global health in a
 538 globalized economy! [Editorial]. *Int J Public Health* 60, 757-759. DOI:
 539 [10.1007/s00038-015-0729-0](https://doi.org/10.1007/s00038-015-0729-0)

540

541

542

543

544

545

546 TABLES AND FIGURES

547

Entire period (25 September 2009 to 21 October 2011)					
Variable	N ^a	Mean	SD ^b	Min	Max
MODIS AOD 3km (unitless)	649	0,29	0,36	0,03	2,19
PM _{2,5} (µg/m ³)	649	11,36	20,06	1,68	164,41
Average temperature (°C)	649	26,82	1,41	16,24	31,26
Relative humidity (%)	649	84,93	5,80	61,50	98,75
Precipitation (mm)	649	5,06	11,37	0	71,40
Dry season (June to November)					
Variable	N	Mean	SD	Min	Max
MODIS AOD 3km (unitless)	323	0,44	0,46	0,03	2,19
PM _{2,5} (µg/m ³)	323	20,51	25,19	1,68	164,41
Average temperature (°C)	323	27,00	1,68	16,24	31,26
Relative humidity (%)	323	82,18	5,90	61,50	98,75
Precipitation (mm)	323	0	0	0	0
Rainy season (December to May)					
Variable	N	Mean	SD	Min	Max
MODIS AOD 3km (unitless)	326	0,14	0,07	0,03	0,76
PM _{2,5} (µg/m ³)	326	2,28	2,87	1,68	26,62
Average temperature (°C)	326	26,66	1,07	22,07	29,05
Relative humidity (%)	326	87,65	4,21	73,80	98,00
Precipitation (mm)	326	7,76	14,09	0,10	71,40
Pearson correlation					
Variable	AOD	PM2.5	TEMP	RH	PRECIP
MODIS AOD 3km (unitless)	1				
PM _{2,5} (µg/m ³)	0.5811	1			
Average temperature (°C)	0.245	0.1007	1		
Relative humidity (%)	-0.3957	-0.4455	-0.4411	1	
Precipitation (mm)	-0.101	-0.1454	-0.1638	0.2413	1

548

Note: a= Number of values observed in days; b=Standard deviation (SD)

549 **Table 1:** Descriptive statistics of the parameters analysed during the study period
 550 (September, 25th 2009 to October, 21th 2011).

551

552

553

554

555

556

557

558

559

Prediction models	N	df ^a	R ²	R ² adj ^b	RMSE ^c
Model 1	649	3	0.544	0.541	13.59
Model 2	649	18	0.683	0.674	11.47
Model 3	649	21	0.707	0.697	11.04
Model 4	649	27	0.782	0.772	9.58
Model 5	649	15	0.823	0.816	8.60

560

Note: a=degrees of freedom; b= R-squared adjusted; c= Residual standard deviation (RMSE).

561

Model 1 = simple model with linear, quadratic and cubic term of AOD;

562

Model 2 = Model 1 + interactions of AOD, AOD² and AOD³ with linear and quadratic terms in temperature and relative humidity;

563

564

Model 3 = Model 2 + interactions between AOD, AOD² and AOD³ and rain;

565

Model 4 = Model 3 + interactions of AOD, AOD² and AOD³ with sine and cosines of date with a period of 365.25 days (without the term for rainy season);

566

Model 5 = Model 4 + lagged relative residual and its square as additional predictor variables;

567

568

569

Table 2: Comparison between prediction models.

570

571

572

Model 1	Estimate	Std. Error	t value	Pr(> t)
β_0	113.565	0.5333	21.295	< 2e-16 ***
β_1	3.733.880	135.860	27.483	< 2e-16 ***
β_2	369.936	135.860	2.723	0.00665 **
β_3	319.761	135.860	2.354	0.01889 *
Model 2	Estimate	Std. Error	t value	Pr(> t)
β_0	1,10E+06	1,89E+05	5.803	1.03e-08 ***
β_1	8,13E+07	9,91E+06	8.202	1.33e-15 ***
β_2	3,19E+07	8,50E+06	3.751	0.000193 ***
β_3	2,32E+06	5,36E+06	0.432	0.665709
β_4	8,85E+04	7,41E+04	1.195	0.232703
β_5	-1,62E+04	1,26E+04	-1.286	0.198907
β_6	-2,44E+03	9,76E+02	-2.503	0.012569 *
β_7	4,02E+02	4,75E+02	0.847	0.397177

β_8	-4,14E+05	2,03E+05	-2.046	0.041196 *
β_9	2,82E+04	3,92E+04	0.720	0.472077
β_{10}	9,39E+03	3,01E+03	3.119	0.001895 **
β_{11}	5,05E+01	1,17E+02	0.431	0.666751
β_{12}	-1,19E+03	1,17E+03	-1.017	0.309673
β_{13}	9,79E+04	9,91E+04	0.988	0.323655
β_{14}	-5,25E+04	2,15E+04	-2.445	0.014744 *
β_{15}	-3,73E+03	1,53E+03	-2.437	0.015070 *
β_{16}	9,53E+01	7,64E+01	1.247	0.212722
β_{17}	1,34E+03	5,54E+02	2.420	0.015795 *
Model 3	Estimate	Std. Error	t value	Pr(> t)
β_0	1,09E+06	1,85E+05	5.870	7.05e-09 ***
β_1	8,22E+07	9,58E+06	8.584	< 2e-16 ***
β_2	3,36E+07	8,29E+06	4.057	5.60e-05 ***
β_3	2,95E+06	5,21E+06	0.566	0.571754
β_4	8,48E+04	7,17E+04	1.183	0.237140
β_5	-1,68E+04	1,22E+04	-1.379	0.168533
β_6	-2,47E+03	9,43E+02	-2.616	0.009109 **
β_7	5,13E+02	4,60E+02	1.116	0.264968
β_8	-4,16E+05	1,96E+05	-2.123	0.034177 *
β_9	4,21E+04	3,87E+04	1.088	0.276841
β_{10}	9,67E+03	2,91E+03	3.325	0.000935 ***
β_{11}	-1,90E+01	1,19E+02	-0.159	0.873546
β_{12}	-1,46E+03	1,14E+03	-1.285	0.199287
β_{13}	8,94E+04	9,59E+04	0.932	0.351465
β_{14}	-5,79E+04	2,12E+04	-2.734	0.006433 **
β_{15}	-3,69E+03	1,48E+03	-2.498	0.012757 *
β_{16}	1,20E+02	7,65E+01	1.572	0.116357
β_{17}	1,46E+03	5,39E+02	2.708	0.006950 **
β_{18}	-1,01E+05	2,00E+04	-5.063	5.44e-07 ***
β_{19}	3,62E+05	1,18E+05	3.063	0.002283 **
β_{20}	-2,58E+05	1,33E+05	-1.935	0.053453 .
Model 4	Estimate	Std. Error	t value	Pr(> t)
β_0	5,55E+05	1,68E+05	3.305	0.001003 **
β_1	3,85E+07	1,04E+07	3.692	0.000242 ***
β_2	4,62E+06	9,84E+06	0.469	0.638882
β_3	-6,61E+06	5,32E+06	-1.245	0.213777
β_4	1,04E+05	6,25E+04	1.659	0.097687 .
β_5	1,28E+02	1,08E+04	0.012	0.990548
β_6	-1,88E+03	8,21E+02	-2.291	0.022301 *
β_7	-1,21E+02	4,08E+02	-0.296	0.767175
β_8	-3,37E+05	1,72E+05	-1.956	0.050898 .
β_9	-5,94E+04	4,05E+04	-1.467	0.142785
β_{10}	4,84E+03	2,56E+03	1.893	0.058828 .
β_{11}	2,10E+02	1,19E+02	1.773	0.076650 .
β_{12}	9,88E+02	1,11E+03	0.894	0.371821
β_{13}	9,74E+04	8,57E+04	1.136	0.256252
β_{14}	2,08E+04	2,51E+04	0.832	0.405926

β_{15}	-1,46E+03	1,30E+03	-1.129	0.259177
β_{16}	-8,98E+01	8,23E+01	-1.091	0.275639
β_{17}	-2,22E+02	5,79E+02	-0.383	0.701887
β_{18}	-1,24E+04	2,54E+04	-0.488	0.625602
β_{19}	7,89E+04	1,19E+05	0.664	0.506904
β_{20}	-9,50E+04	1,24E+05	-0.767	0.443244
β_{21}	6,99E+04	1,49E+04	4.681	3.51e-06 ***
β_{22}	-3,71E+04	7,47E+03	-4.966	8.83e-07 ***
β_{23}	-2,64E+05	5,20E+04	-5.068	5.31e-07 ***
β_{24}	4,90E+04	2,17E+04	2.257	0.024378 *
β_{25}	1,08E+05	3,06E+04	3.545	0.000423 ***
β_{26}	-3,07E+04	1,37E+04	-2.244	0.025199 *
Model 5	Estimate	Std. Error	t value	Pr(> t)
β_0	7,46E+03	1,19E+03	6.266	6.92e-10 ***
β_1	1,18E+06	1,62E+06	0.732	0.464505
β_2	4,13E+06	3,97E+06	1.041	0.298197
β_3	-1,48E+06	1,97E+06	-0.754	0.451224
β_4	5,84E+04	6,22E+04	0.938	0.348512
β_5	-4,32E+04	2,33E+04	-1.858	0.063575 .
β_6	-1,53E+03	7,45E+02	-2.049	0.040855 *
β_7	2,09E+02	8,55E+01	2.443	0.014849 *
β_8	2,10E+02	4,51E+02	0.465	0.641789
β_9	-3,03E+05	1,66E+05	-1.833	0.067247 .
β_{10}	5,14E+01	5,64E+04	0.001	0.999274
β_{11}	4,05E+03	2,31E+03	1.750	0.080635 .
β_{12}	-1,70E+02	2,04E+02	-0.833	0.404931
β_{13}	1,11E+03	1,13E+03	0.981	0.327124
β_{14}	9,58E+04	8,07E+04	1.187	0.235520
β_{15}	2,21E+03	2,94E+04	0.075	0.939958
β_{16}	-1,14E+03	1,17E+03	-0.975	0.329859
β_{17}	5,66E+01	1,06E+02	0.533	0.594386
β_{18}	-4,24E+02	5,72E+02	-0.741	0.459242
β_{19}	7,01E+04	7,96E+03	8.812	< 2e-16 ***
β_{20}	-3,67E+04	5,74E+03	-6.389	3.27e-10 ***
β_{21}	-2,55E+05	3,53E+04	-7.223	1.49e-12 ***
β_{22}	5,66E+04	1,77E+04	3.195	0.001470 **
β_{23}	1,18E+05	2,26E+04	5.234	2.27e-07 ***
β_{24}	-4,10E+04	1,17E+04	-3.504	0.000491 ***
β_{25}	4,55E+02	3,75E+01	12.129	< 2e-16 ***
Model 6	Estimate	Std. Error	t value	Pr(> t)
β_0	2,35E+03	9,58E+01	24.568	< 2e-16 ***
β_1	-1,12E+04	5,55E+03	-2.019	0.0439 *
β_2	2,02E+04	2,16E+03	9.358	< 2e-16 ***
β_3	-9,94E+03	9,56E+02	-10.405	< 2e-16 ***
β_4	1,71E+02	3,39E+02	0.503	0.6154
β_5	5,82E+01	7,62E+01	0.764	0.4454
β_6	-3,13E+00	6,32E+00	-0.495	0.6211
β_7	-4,23E-01	4,83E-01	-0.875	0.3817

β_8	8,42E+03	9,67E+02	8.708	$< 2e-16$ ***
β_9	-6,59E+03	4,36E+02	-15.117	$< 2e-16$ ***
β_{10}	-2,08E+04	1,97E+03	-10.561	$< 2e-16$ ***
β_{11}	1,01E+04	8,26E+02	12.217	$< 2e-16$ ***
β_{12}	1,00E+04	8,98E+02	11.130	$< 2e-16$ ***
β_{13}	-4,31E+03	3,51E+02	-12.275	$< 2e-16$ ***
β_{14}	5,57E+02	4,34E+01	12.841	$< 2e-16$ ***
β_{15}	-8,85E+01	1,11E+01	-8.012	5.45e-15 ***

Note: Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

573

574 **Table 3:** Description of parameters, standard error and p-value for each prediction
575 models.

576

577

578

579

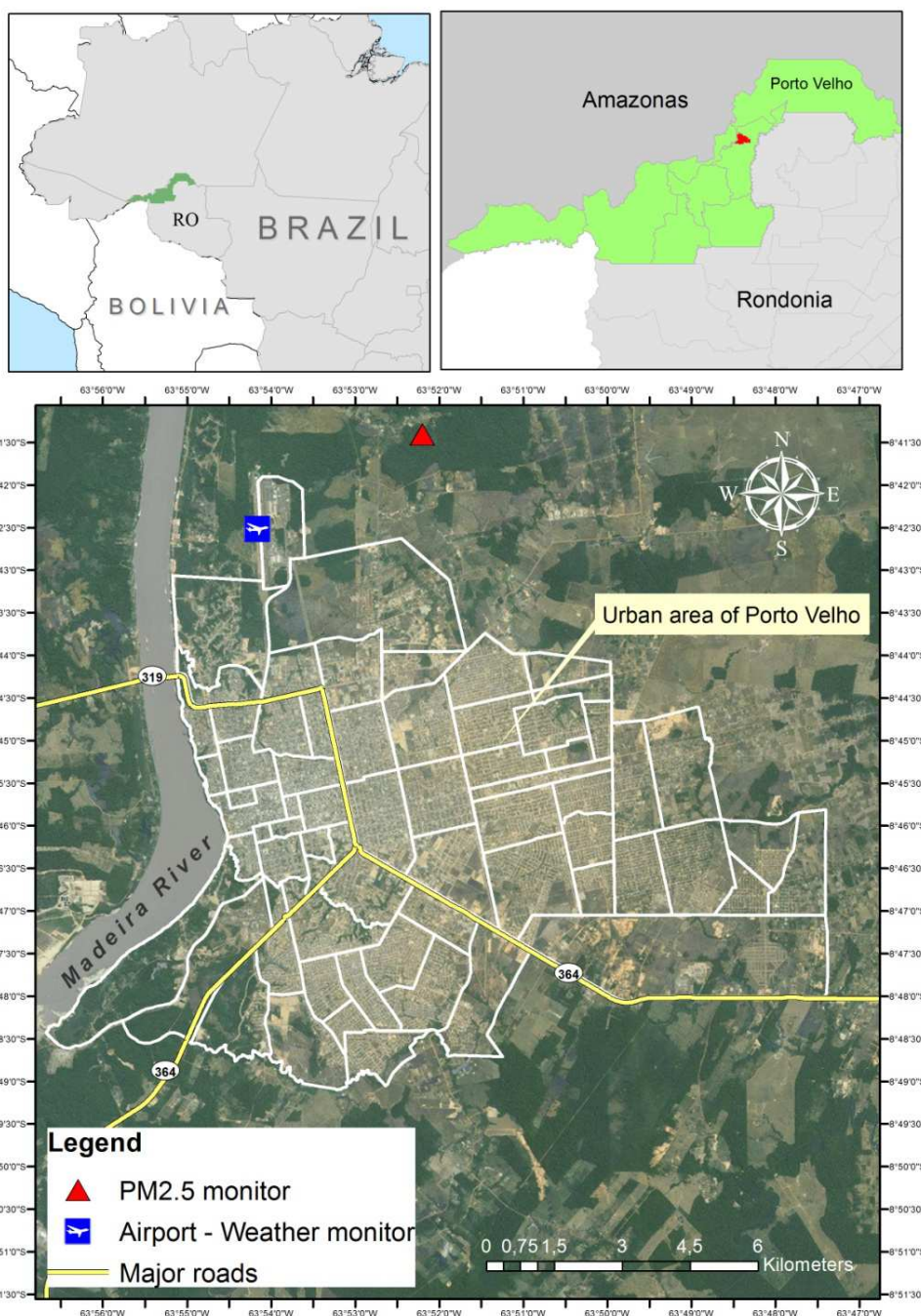
580

581

582

583

584



600

601

602 **Figure 1:** Study area with the locations of PM_{2.5} and weather monitors. Municipality
603 of Porto Velho, Rondônia state, Brazilian Amazon region.

604

605

606

607

608

609

610

611

612

613

614

615

616

617

618

619

620

621

622

623

624

625

626

627

628

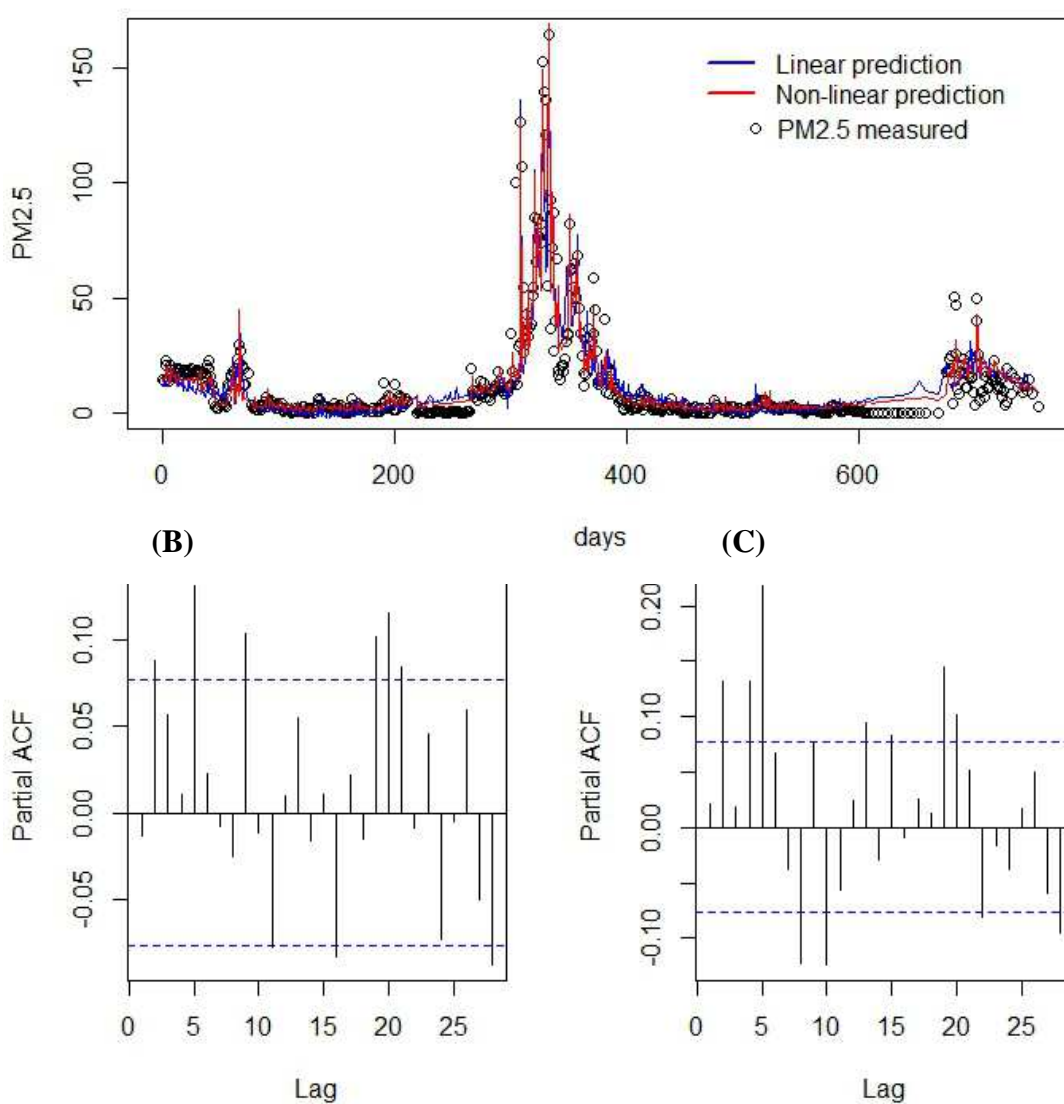
629

630

631

632

633



634 **Figure 2:** (A) Comparisons between PM_{2.5} measured and PM_{2.5} predictions ($\mu\text{g}/\text{m}^3$)
635 across time (25 September 2009 to 21 October 2011). (B) Partial autocorrelation plot of
636 residuals before introducing the lagged relative residual and its square as additional
637 predictor variables into the model (C) Partial residual autocorrelation plot after
638 adding the the two variables to the model.

639

640

641

642

643

644

645

646

647

648 PM2.5 predicted

649

650

651

652

653

654

655

656

657

658

659

660

661

662

663

664

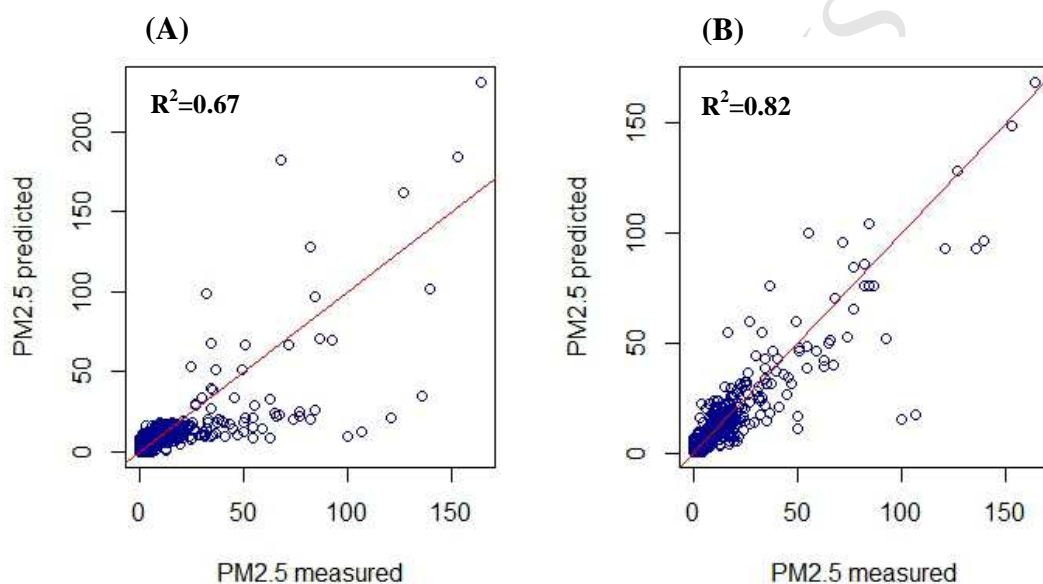
665

666

667

668

669



660 **Figure 3:** Comparisons between the measured and predicted PM_{2.5}($\mu\text{g}/\text{m}^3$) for (A)
661 Model 2 and (B) non-linear prediction (Final model). The red line represents the
662 regression line.

670

671

672

673

674

675

676

677

678

679

680

681

682

683

684

685

686

687

688

689

690

691

692

693

694

695

696

697

698

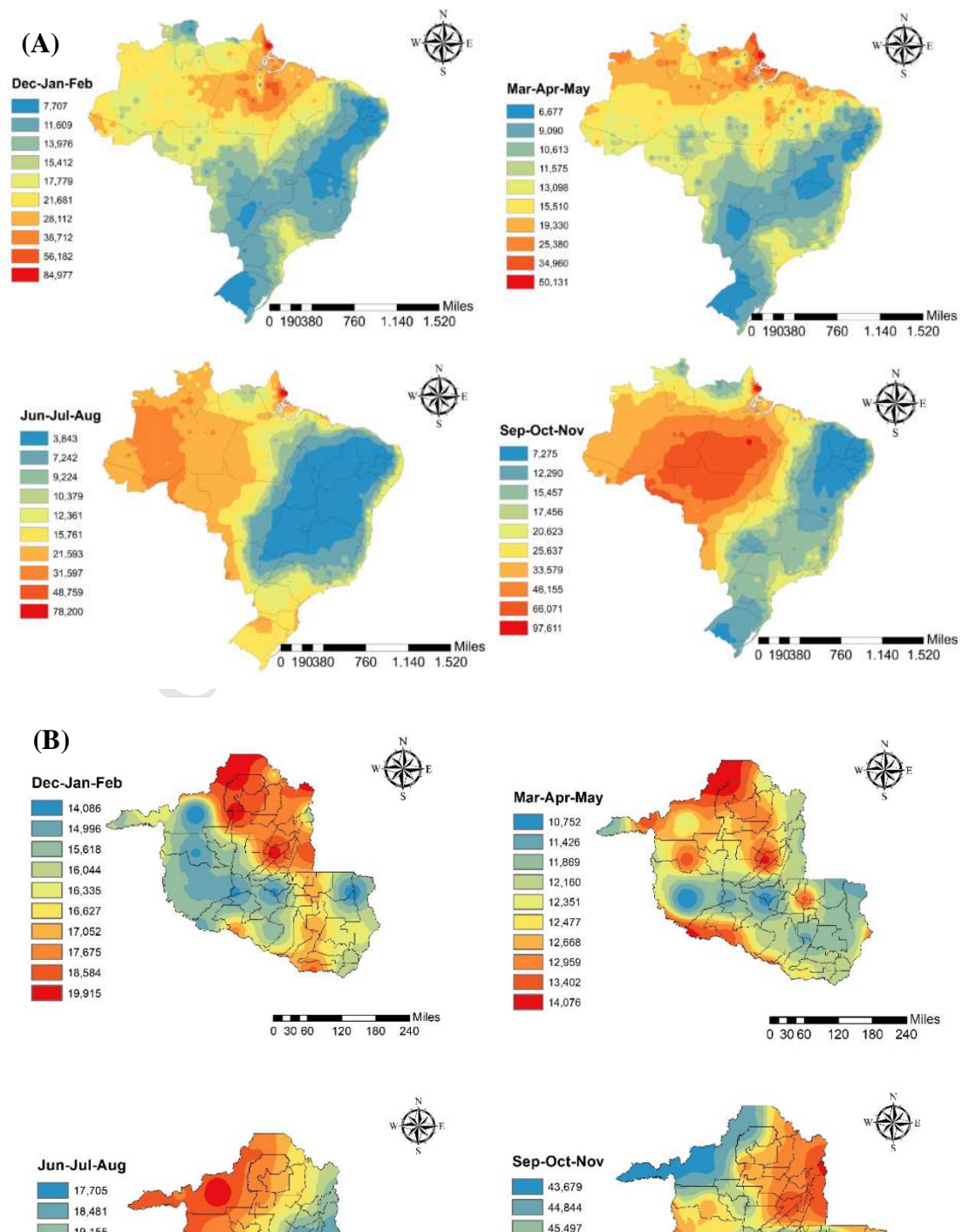
699

700

701

702

703



704

705

706

707

708

709

710

711 **Figure 4:** Spatial distribution of PM_{2.5} predicted concentration over the basin during
712 different seasons. Forest fires occurs between the months of September, October and
713 November. (A): PM_{2.5} averages predicted concentration interpolated to all Brazilian
714 states; (B): Rondônia State including the study area of Porto Velho.

Highlights

- Non-linear model was applied for predicting $PM_{2.5}$ from AOD with a good performance;
- The model can be applied to other sites if site-specific data are available;
- The lagged relative residual it is cautious strategy to further improve the model;
- It was the first Brazilian study predicting $PM_{2.5}$ from high resolution AOD data;